

X-KEYSCORE as a SIGDEV tool

2009

## What is X-KEYSCORE?

#### What is XKEYSCORE?



# A (DNI) SIGDEV Tool

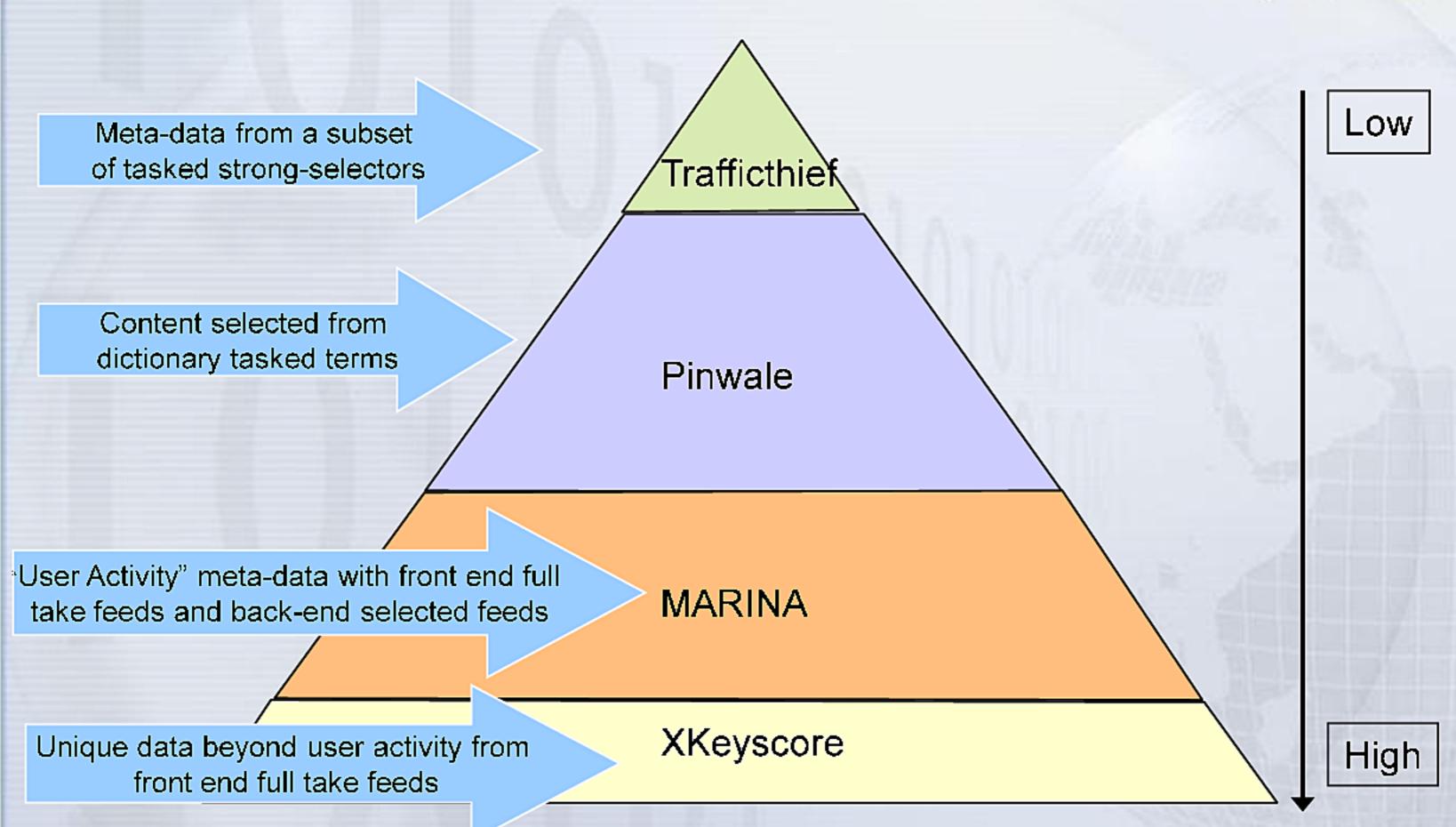
It gives you the ability to discover things that you otherwise wouldn't have seen

#### What makes XKS so good at SIGDEV?

- 23 CESCORE
- XKS gives analysts unique access to terabytes of content and meta-data
- Typically sites select and forward to PINWALE less than 5% of the DNI they're processing
- The rest of that data used to be dropped but is now being retained temporarily and made available to analysts through X-KEYSCORE
- As an example, at one our sites XKS sees more data per day than all of PINWALE

# DNI Discovery Options





#### "Slowing down the Internet"



- XKS goal is to store the full-take content for 3-5 days, effectively "slowing down the Internet" so that analysts can go back and recover sessions that otherwise would have been dropped by the front end
- Meta-data is saved off longer, with the goal of 30 days retention
- A lot of analysis can be done through meta-data only (MARINA is meta-data only)





- Front end storage is limited by resources and policy restrictions and will vary by site
- At some sites, the amount of data we receive per day (20+ Terabytes) can only be stored for as little as 24 hours based on available resources
- Other sites have legal or policy restrictions that limit the amount of time we can store data (if we can at all)
- It's a rolling buffer where new data comes in and pushes the oldest data out



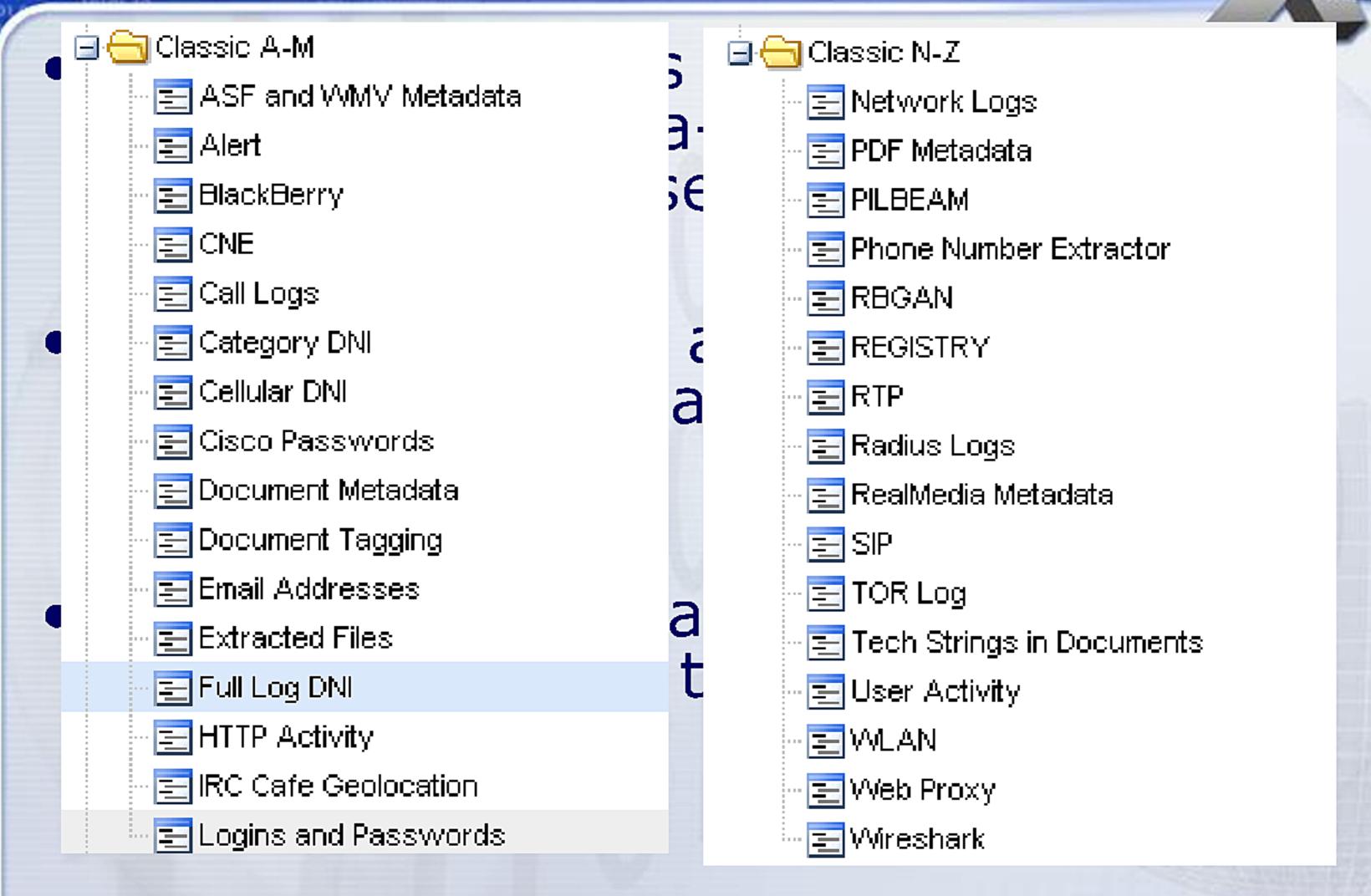


- Content that is "interesting" can be pulled out of X-KEYSCORE and pushed to Agility or PINWALE or any other database for longer retention
- Workflows can be set up to automatically harvest content out of XKS before it ages off
- The goal, however, is to use X-KEYSCORE to discover new things, that will end up on tasking for future collection

#### How do I access XKS data?

- 23 CESSCORE
- It's important to know that XKS queries meta-data tables only
- Results from the meta-data tables are then linked back to the original piece of content
- Goal of the system is to extract a wide range of meta-data for users to query

## What kind of meta-data is produced?



### Examples of "simple" Plug-ins



Plug-in	DESCRIPTION
E-mail Addresses	Indexes every E-mail address seen in a session by both username and domain
Extracted Files	Indexes every file seen in a session by both filename and extension
Full Log	Indexes every DNI session collected. Data is indexed by the standard N-tupple (IP, Port, Casenotation etc.)
HTTP Parser	Indexes the client-side HTTP traffic (examples to follow)
Phone Number	Indexes every phone number seen in a session (e.g. address book entries or signature block)

## Examples of "advanced" Plug-ins



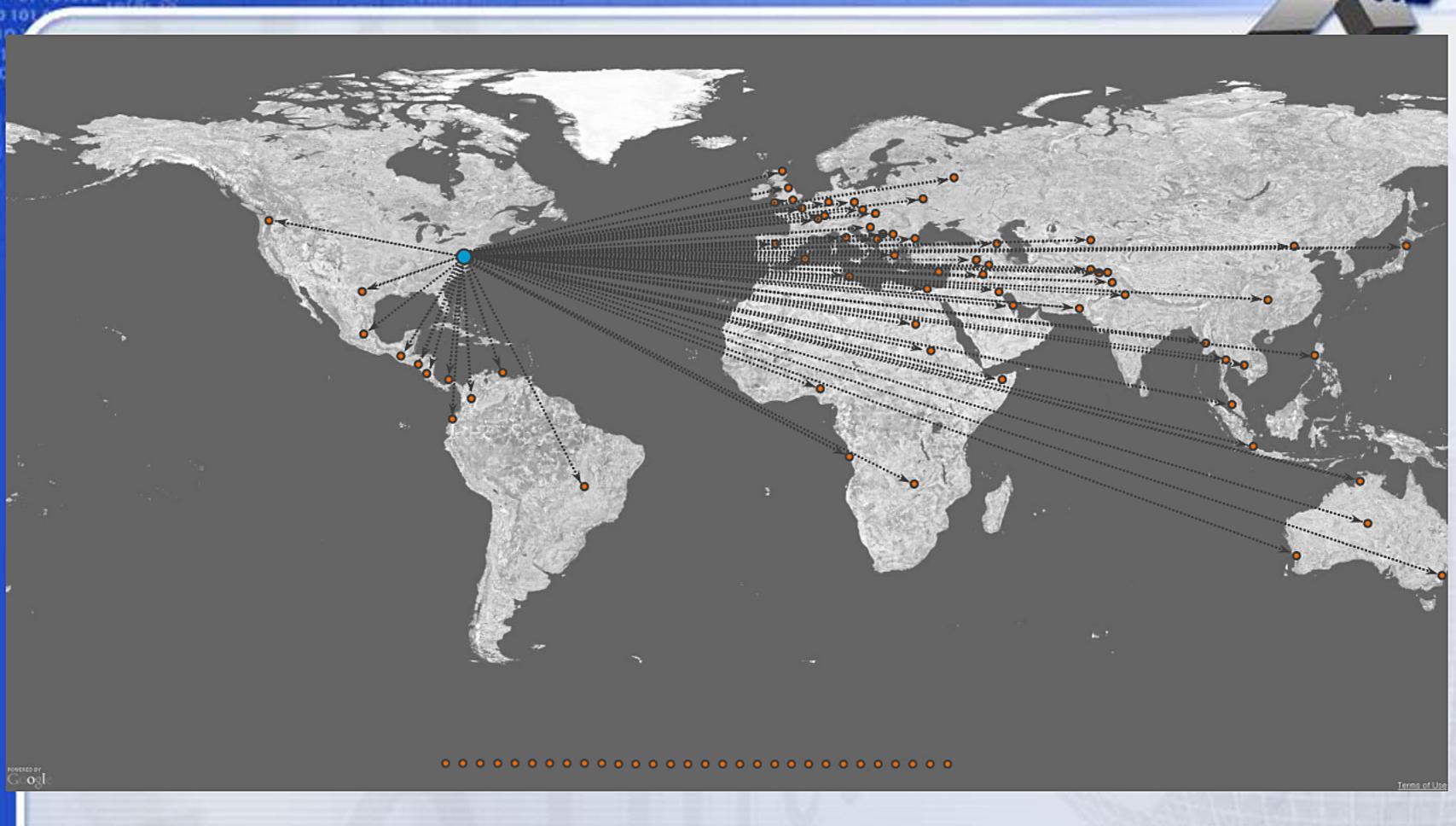
Plug-in	DESCRIPTION
User Activity	Indexes the Webmail and Chat activity to include username, buddylist, machine specific cookies etc. (AppProc does the exploitation)
Document meta- data	Extracts embedded properties of Microsoft Office and Adobe PDF files, such as Author, Organization, date created etc.

#### Plug-ins

- A single session may contain entries in multiple meta-data tables
- For example, if a single session had a user Emailing an attached word document the following plug-ins would extract meta-data:

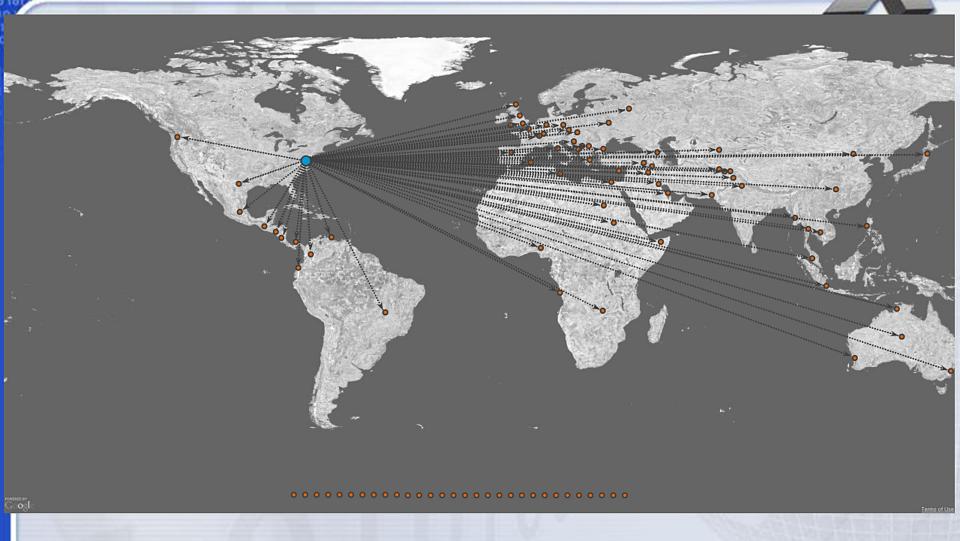
Plug-in	Would have extracted
Full Log	bare minimum meta-data like To/From IP address, ports, casenotation, sigad etc.
E-mail Addresses	any E-mail addresses seen on that page (including inside the attached word file)
Extracted Files	the filename and extension of the attachment
Document Meta- data	in addition to the filename and extension, any embedded properties of the word document like Author, last author, organization, date created, date last modified etc.

#### X-KEYSCORE DEPLOYMENTS



125 Sites

#### X-KEYSCORE DEPLOYMENTS



125 Sites

TOP SECRET//COMINT//ORCON, REL TO USA, AUS, CAN, GBR and NZL//20291123

# AppIds and Fingerprints



- X-KEYSCORE produces an application id for each session processed
- Currently almost 1300 Appids in 28 categories
- An Appid is meant to identify a session as a particular application
- Fingerprints are an extensible way of tagging sessions
- Ex: A session Appid'd as mail/smtp might also contain fingerprints for encryption if used in the email

# AppIds and Fingerprints



#### Ex: E-Mails with encryption

From: "Launchpad OpenPGP Key Confirmation" <noreply@launchpad.net> [Save Address] [Block Sender]

To:

Cc:

Subject: Launchpad: Confirm your OpenPGP Key

Date: Wed, 31 Dec 2008 10:04:16 -0000

----BEGIN PGP MESSAGE----

Version: GnuPG v1 4 6 (GNU/Linux)

Application

AppID (+Fingerprints)

mail/webmail/outblaze

mail/webmail/outblaze has\_fingerprint encryption/pgp encryption/pgp/message

spflvtVPZsl1vpg67VdHFUprgv0JpmjQlb73gWmhbOUrZzyGdDRla9CcFzJA7OIL 3XyCrlniniJ4/c98+khDazh1XY/S7yNi38Wrlkd3GOz9DFFI1Nu31nwjh3+nc0pv OlyztsQzLFB/8+qJrPvmK8fzz7tWp2djxyfMGoAYNAf/QOohR0BjqTgOUlqLRVrE eEFivrMOnBxf60SHIFra7LpZlsTUFpBJNAkgguk7m8fJ0dMmU0V5MeM1x8GuVW5+i Uk4bBwwZ1VpEVHCyGuv8ux+V+KpSkQtDwdhlp12SZ2SUm1upnVB9lfcnlhVVxZp LaY3mXqNWhyhzFPFxkhUwqzd/rMxrCJucfXGaeisSizZDIQOWxTSwe7BwvG8Bvnr QEQVKY30vWg+2pDTPrKq3uEqOwj9JY7KTPMrf2gZLNABDuCJm5lRALZqqETTg4dhi xV0r9+2ZLtyGDXQhLMyBElYns4+jiP1rd3E+TW7JVUe/dPluyC4DwOUPklwuHcC+ StLAuQHMS6RkB4aDNdi6QG9kEVWjq2PvfuMIBWo5jJ8RFoDSx8q5t1ukgeCxr6xr Q4eTmOFTIA71G312Xa7ZniOzyxiWZ4CAbhHLF+3baFD3lb4/EFmRvPBdqy6wUyHD Z5EXyHDzI4XIDyEe/aomEqAsUqPs8MZirHHzpbaS3LbG5B5VKAKU59bENpf/K0gTi a3IUAeQ1t6xLzgToVdfhEkPj5bxODrWcZtHeTEt1nV+3pc2P58+QICDOETiDCA/j. dhG2brUwbxny6Ap7fU5e1ALU3ryoXKvt9eCXZHooY/p9QlC3koHCWptGD6gKCxlt KWWK5M+HkxhHy4V7Wb137CStzeLda8BdU43Kh0ZQWWjK7pDXKKhHLYlGlawRScQal e6J+y4JR1KKyXiXY94Erxa/P0FzuYV/QCJUDpqWFR22bXuy4FhkosLWM8G+UBHVt UfgRxq8as60DhBDWy08eLEAdE92TVffJqXOvAOzTqBrP7uZi/Q7ABFFGTQ9ni =N4CJ

----END PGP MESSAGE----

TOP SEC Thanks,

10291123

# AppIds and Fingerprints



#### **Ex: Airline E-Tickets**

Subject: Airblue E-Ticket - JGDTGSWB

From: Airblue Reservations <website@airblue.com>;

To:

**Date:** Nov 18, 2008 10:41:54 AM

#### airblue

Reservation No.

Application.

AppID (+Fingerprints)

mail/webmail/yahoo

mail/webmail/yahoo has\_fingerprint travel/airblue.

_			•	
Phone 1: Phone 2:		AZ2749951 21 6876 5648 15	E	
Travel Coupon(s	): PHOTO	ID REQUIRED AT CHECK-IN		
21 6876 564	8/1 ED 610	Peshawar Dubai 29 NOV 08 12:0	0 O YA90O	15,505.00 OK
Transactions hist	ory details:			
Date/Time	Method	Location	Description	Amount
18-Nov-2008 3:41 PM	Travel Agency	Khaleej Express- Pew Muhammed Younas, Main Branch	Ticket Sale	Rs 15,505.00

TOP SECRET//COMINT//ORCON, REL TO USA, AUS, CAN, GBR and NZL//20291123

# AppIds and Fingerprints



■ HTTP Header Information

Content Type: HTTP/POST/Form-Data

POST Ab/private.php?do=insertpm &pmid= HTTP/1.1

image/gif, image/x-xbitmap, image/jpeg, image/pjpeg, application/x-shockwave-flash, application/vnd.ms-excel, Accept:

application/vnd.ms-powerpoint, application/msword, \*/\*

Referer:

Accept-Language: en-gb

Content-Type: application/x-www-form-urlencoded

UA-CPU: x86

Accept-Encoding: gzip, deflate

User-Agent: Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.1; FDM)

AppID (+Fingerprints) Application.

mail/webmail/vbulletin/private\_message/insert has\_fingerprint forum/extremist/al-faloja mail/webmail/vbulletin/private\_message/insert

recipients	
bccrecipients	
title	خبر مهم
	شنت كتبية المواجهات التابعة لحركة الشباب المجاهدين-بفضل الله مساء بوم الإنتين 08 محرم 1430هـ الموافق لـ:05-01-2009م هجوما مباشرا وعنيفا على مصنح الباسئا للغوات الصلبيبة الإنبوبية في مغيشو ويشاركت كتبية المدفعيات في العملية المباركة حبت قامت بفصف المصنح بوابل من الصواريخ والمدفعيات
message	
	واستخدم المجاهدون في الهجوم أساليب فتالية غير مسبوقة مما أرغم على قوات العدو التراجع من دفاعاتها في الشارع العام المؤدي إلى المصنع، وحينما

TOP SECRET//COMINT//ORCON, REL TO USA, AUS, CAN, GBR and NZL//20291123

اجتمعوا على جحرهم فاجأتهم كتببة المنفعيات بفصف عنيف وبفوق ويتوقع خسائر بشرية جسيمة في صفوف القوات الصليبية ولله الحمد والمنة

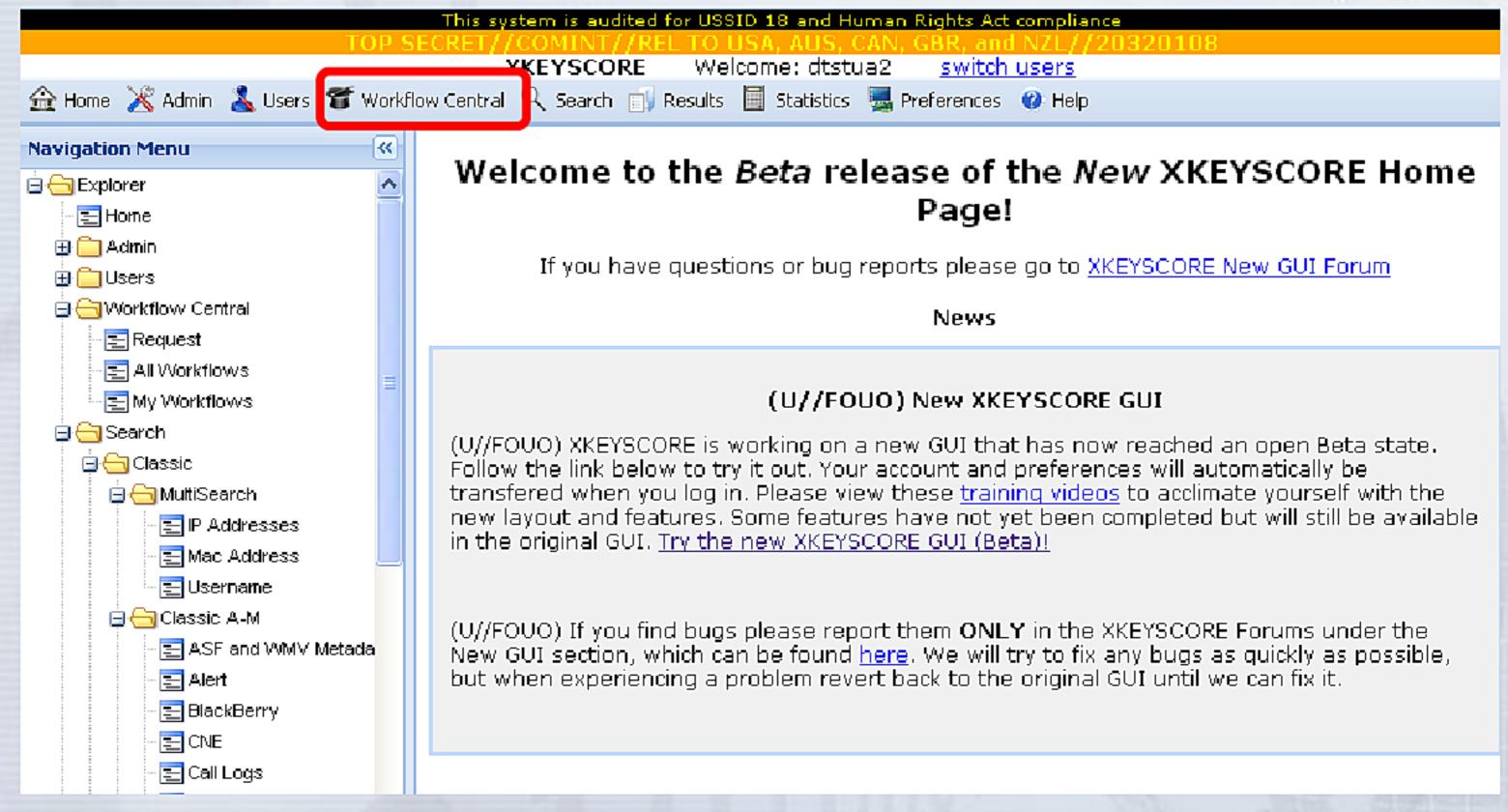
#### X-KEYSCORE Workflows



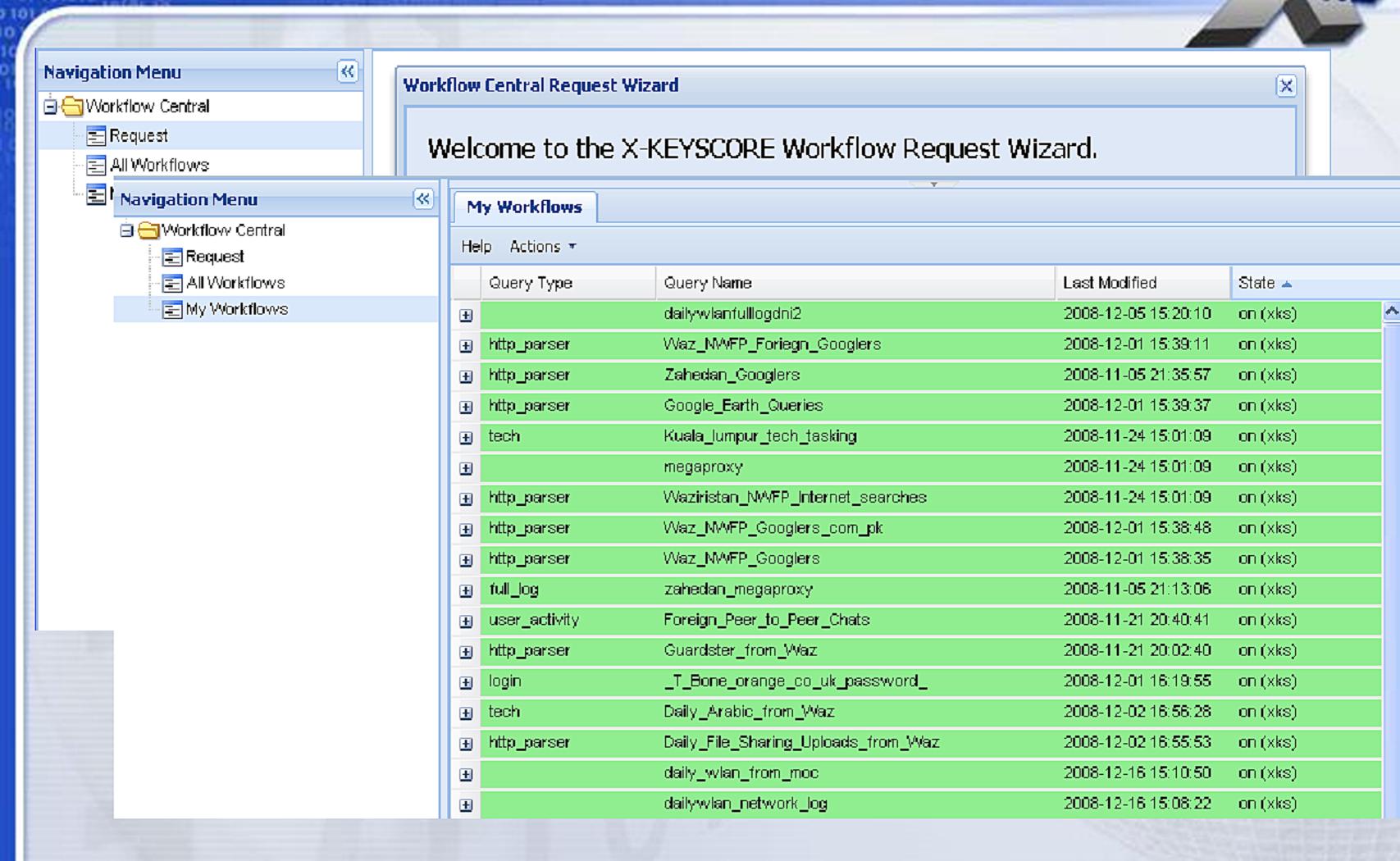
- X-KEYSCORE workflows are standing queries that run on set intervals during the day (usually once a day)
- After action reports can E-mail the results of the workflow, parse out data to mailorder to other databases and more
- New GUI's Workflow Central makes it easy to create and manage your workflows

#### XKS Workflows: Easy to Create!





#### XKS Workflows: Easy to Create!



## Context-Aware Tagging

- 23 CENSIONE
- Provides for the ability to task and scan for terms only when they appear inside the body of documents like Microsoft Office or Adobe PDFs
- EX: We want to find technical documents regarding WIMAX networks but tasking the term 'WIMAX' to Cadence would flood PINWALE with hits. What if we only look for the term within documents?

TOP SECRET//COMINT//ORCON, REL TO USA, AUS, CAN, GBR and NZL//20291123 Y/99 WATEEN TECH TECH DATETIME DATETIME TECH VALUE ID **END** NAME FILENAME NIB Ranchor 2008-01-01 2008-01-01 wireless WIMAX 04:55:00 04:55:01 Line KHI.doc NIB Ranchor 2008-01-01 2008-01-01 2 satellite DVB 04:55:00 04:55:01 Line KHI.doc NIB Ranchor 2008-01-01 2008-01-01 3 mac 04:55:00 04:55:01 Line KHI.doc 7 BUC Make 8 BUC Frequency 9 LNB Type. Κu 1 LNB Frequency

TOP SECRET//COMINT//ORCON, REL TO USA, AUS, CAN, GBR and NZL//20291123

DVB STM 1000

Serial

1 DVB-RCS Modern type

DVB-RCS Modern Serial

# Context-Aware Tagging





TOP SECRET//COMINT//ORCON, REL TO USA, AUS, CAN, GBR and NZL//20291123

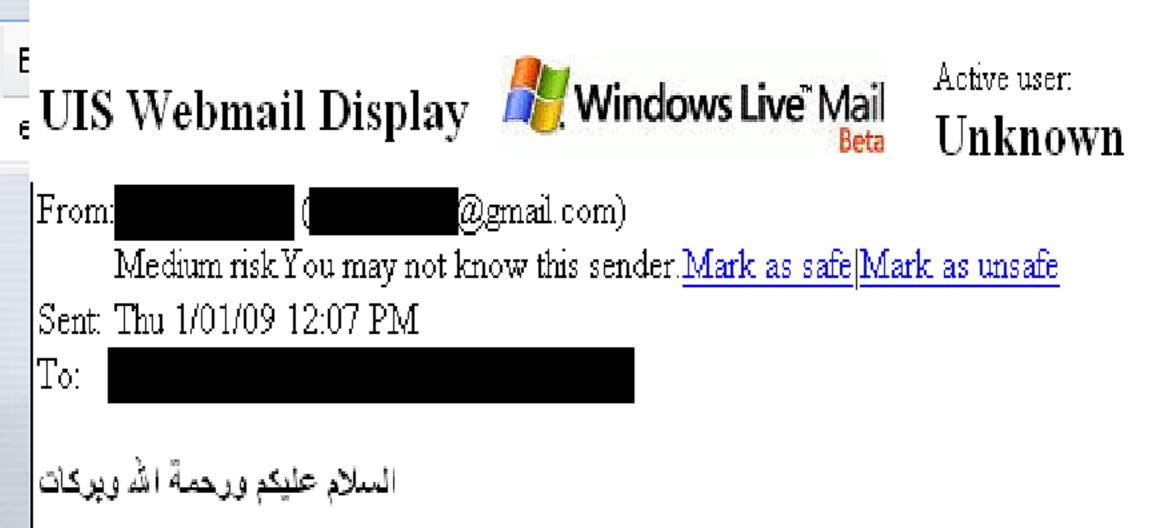
# Context-Aware Scanning



- Tasking is so flexible that it can include regular expressions (REGEXs) with few or no anchor points
- Ex: Can we find documents that have MAC addresses in them?
- The following Regex looks for MAC addresses:
- •"(00|01|02|04|08|10|3C|44):(?=[\d:]{0,1 2}[a-f])([\da-f]{2}):([\da-f]{2}):([\daf]{2}):([\da-f]{2}):([\da-f]{2})"

# Context-Aware Scanning

- 23 CENSIONE
- Supports full foreign language tagging and querying
- •Ex look for common Arabic expressions in E-mails coming from the Pakistan tribal regions:



### X-KEYSCORE SIGDEV



- X-KEYSCORE's full take database of meta-data and content make it an powerful SIGDEV tool
- Many DNI applications don't contain strong selectors that allow traffic to be collected
  - Web surfing
  - Internet searching
  - Anonymous file uploading/downloading
- The variety of applications processed and meta-data available make X-KEYSCORE an ideal starting point for DNI development

### X-KEYSCORE SIGDEV



- Scenario 1: Persona Analysis
- Goal to identify the "user session"
- Help answer the question: What did my target do while he was online?

 We may know from TRAFFICTHIEF, PINWALE or MARINA that our target was online at a given time and from a given IP address, so we can then search in X-KEYSCORE for everything that happened "around" that event.

# XKS SIGDEV: Persona Analysis

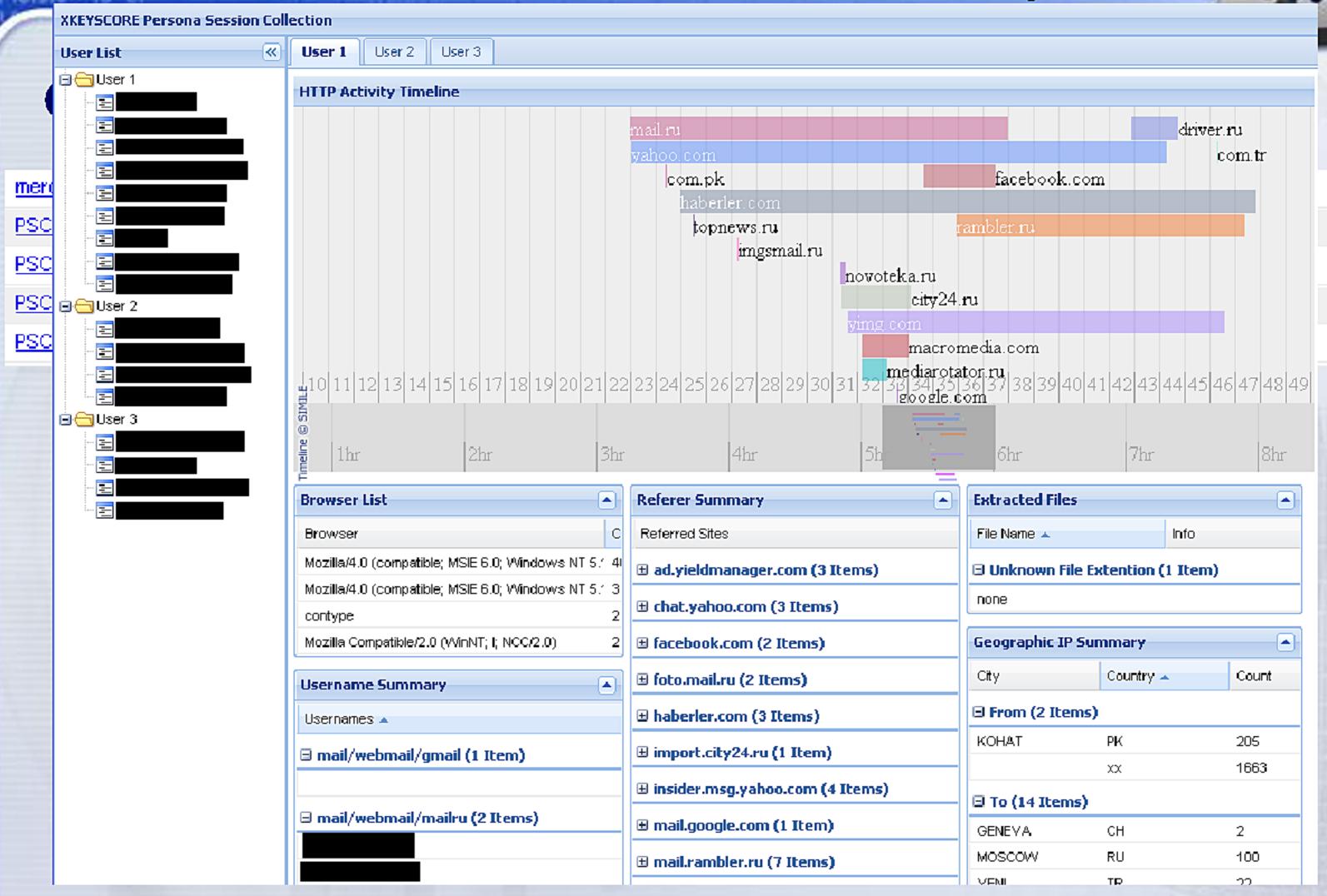
TS ▲	ACTIVE_USER	ACTIVE_USER_I	P ACTI			
20081229 051406Z	≤yahoo>	119.	PK			
20081229 051406Z ;	Datetime 🔺	Search For	Datetime End	Search Value	Fm IP	To IP
20081229 051406Z :	2008-12-29 05:14:07	username	2008-12-29 05:14:18	@yahoo	119.	209.
	2008-12-29 05:14:07	username	2008-12-29 05:14:18	@yahoo	119.	209.
20081229 051407Z :	2008-12-29 05:14:07	username	2008-12-29 05:14:18	@yahoo	119.	209.
20081229 051409Z :	2008-12-29 05:14:07	username	2008-12-29 05:14:18	@yahoo	119.	209.
20081229 051410Z ;	2008-12-29 05:14:07	username	2008-12-29 05:14:18	@yahoo	119.	209.
	2008-12-29 05:14:07	username	2008-12-29 05:14:18	@yahoo	119.	209.
20081229 051410Z ;	2008-12-29 05:14:07	username	2008-12-29 05:14:18	@yahoo	119.	209.
20081229 051410Z :	2008-12-29 05:14:09	username	2008-12-29 05:14:21	@yahoo	119.	209.
20081229 051411Z :	2008-12-29 05:14:09	username	2008-12-29 05:14:21	@yahoo	119.	209.
20081229 051414Z ;	2008-12-29 05:14:09	username	2008-12-29 05:14:21	@yahoo	119.	209.
	2008-12-29 05:14:09	username	2008-12-29 05:14:21	@yahoo	119.	209.
20081229 051415Z :	2008-12-29 05:14:09	username	2008-12-29 05:14:21	@yahoo	119.	209.
20081229 051420Z :	2008-12-29 05:14:09	username	2008-12-29 05:14:21	@yahoo	119.1	209.
20081229 051420Z :	2008-12-29 05:14:09	username	2008-12-29 05:14:21	@yahoo	119.	209.
	2008-12-29 05:14:10	username	2008-12-29 05:14:50	@yahoo	119.	209.
20081229 051421Z :	2008-12-29 05:14:10	username	2008-12-29 05:14:50	@yahoo	119.	209.
20081229 051426Z :	2008-12-29 05:14:10	username	2008-12-29 05:14:50	@yahoo	119.	209.
	2008-12-29 05:14:10	username	2008-12-29 05:14:50	@yahoo	119.	209.
	2008-12-29 05:14:10	username	2008-12-29 05:14:50	@yahoo	119.	209.
	2008-12-29 05:14:10	username	2008-12-29 05:14:50	@yahoo	119.	209.
	2008-12-29 05:14:10	username	2008-12-29 05:14:50	@yahoo		

# XKS SIGDEV: Persona Analysis

#### Coming soon: XKS PSC query builder/viewer

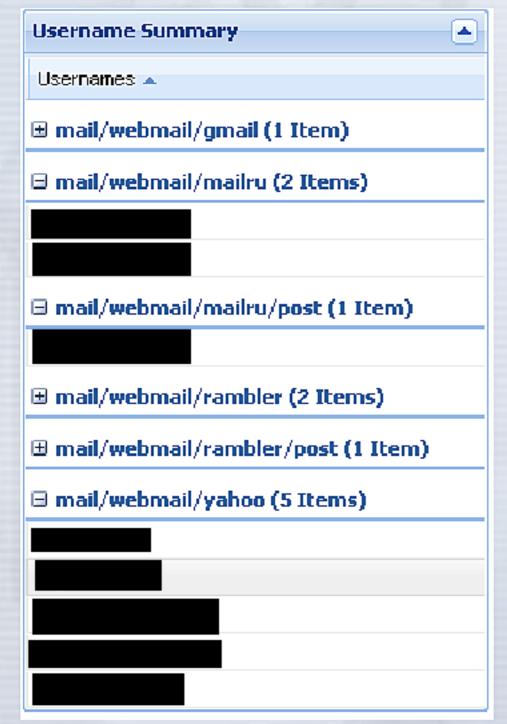
usernam	e 2008-12-29 05:14:	18 @yahoo	user_re	
usernam	e 2008-12-29 05:14:	18 Row Actions	Вс	
usernam		24	nı	
usernar <sup>2</sup>	Persona Session Collectio		×	
usernar 2				٨
usernar <sub>2</sub>	Justification:	Persona session collection for to_ip = 209.191.120.3(		
usernar <sub>2</sub>	Additional Justification:			
usernar	Start Date & Time:	12/29/2008 05:09 (M/D/Y H:M)		
usernar				
usernar	Stop Date & Time:	12/29/2008 🖭 05:19 💌		
usernar	IP (Country Code):	119. (pk)		
usernar				
usernar	Also Query IP As:	✓ From		
usernar <sup>2</sup>		▼ To		
usernar2		X-Forwarded-For IP		
Hearnar	XFF or Client IP:			
2	Add Search:	Extracted File		

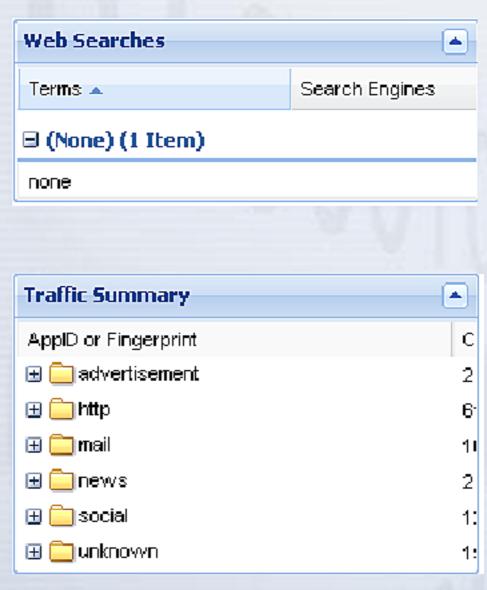
# XKS SIGDEV: Persona Analysis



# XKS SIGDEV: Persona Analysis

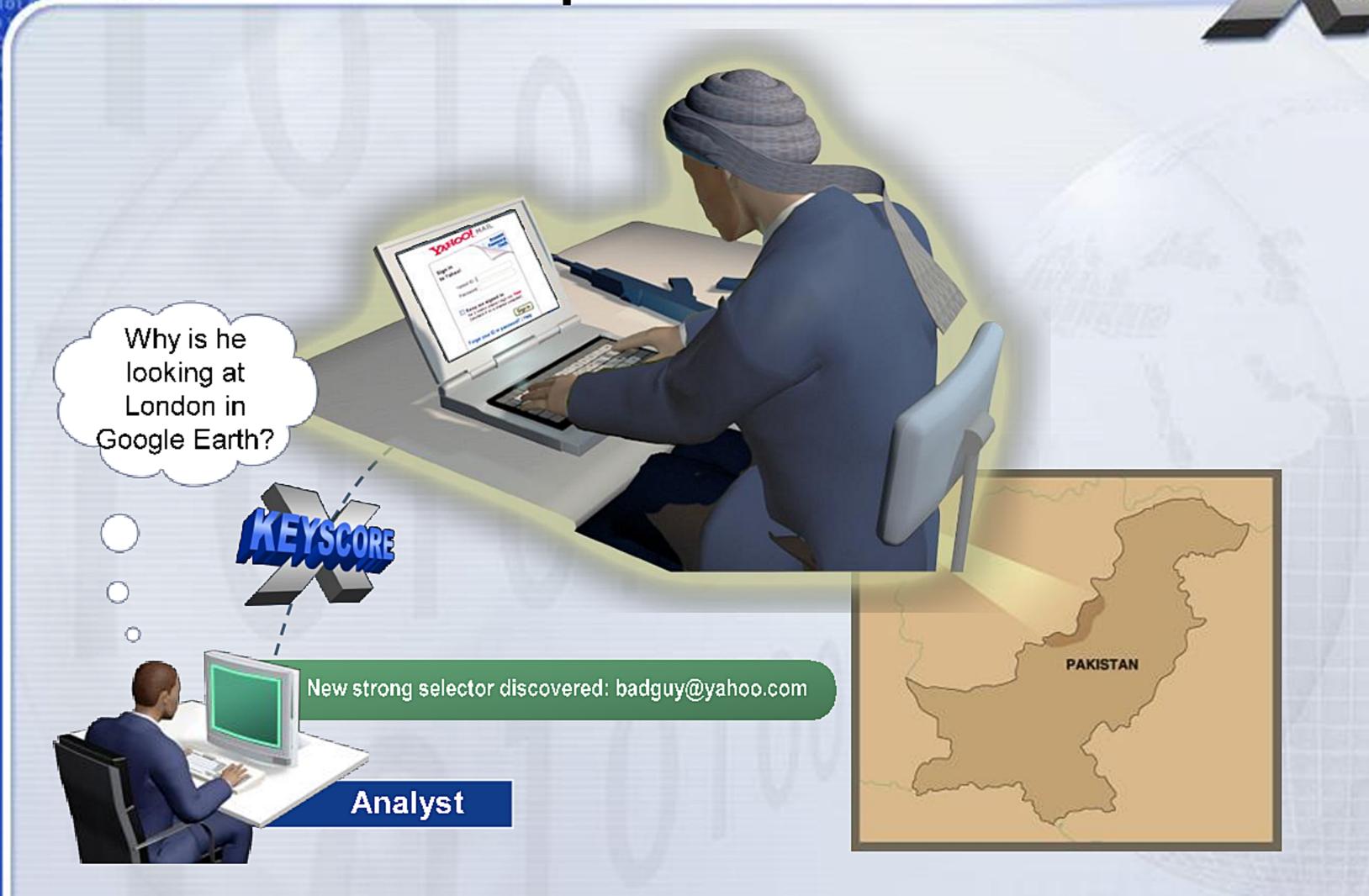
#### Coming soon: XKS PSC query builder/viewer





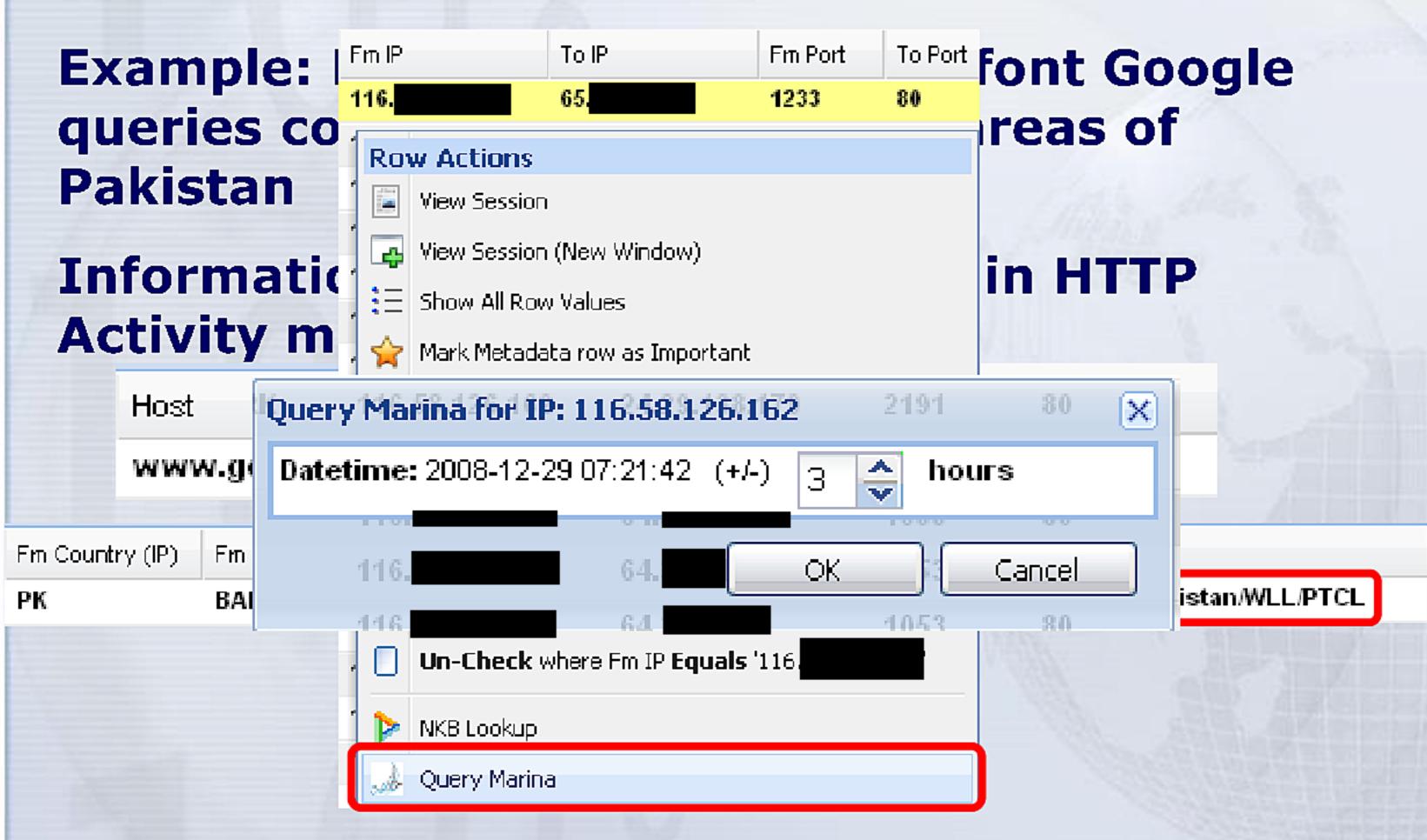
Domain Summary
Subdomains 🔺
⊕ adinterax.com (2 Items)
⊞ adriver.ru (1 Item)
⊞ akamai.net (1 Item)
⊎ bn5.ru (1 Item)
⊞ city24.ru (1 Item)
⊞ com.pk (1 Item)
⊞ com.tr (1 Item)
⊞ facebook.com (2 Items)
⊞ fbcdn.net (1 Item)
⊞ gismeteo.ru (1 Item)
⊞ google.com (1 Item)
⊞ haberler.com (3 Items)
⊎ imgsmail.ru (1 Item)
⊞ macromedia.com (3 Items)
⊞ mail.ru (7 Items)

# NWFP Example



#### XKS SIGDEV: HTTP Traffic





### XKS SIGDEV: HTTP Traffic



#### XKS SIGDEV: HTTP Traffic



#### Now make that into a workflow

```
X-KEYSCORE EMAILER
QUERY NAME: Waz NWFP Foriegn Googlers
 current time: 2008-11-20 07:15:15 GMT
 submitted at: 2008-11-20 03:55:03 GMT
 has 14 result(s)
www.google.com
                2008-11-19 18:54:20 : al qaida (en, en-GB) (1)
 116.
                2008-11-19 07:36:49 : The al-Ikhlas network (cybertrans from Arabic) (1)
 116.
 116.
                2008-11-19 07:37:07 : (referer) the al-Ikhlas network (cybertrans from Arabic) (3)
 116.
                2008-11-19 08:03:17 : Forum bride/'Arus (cybertrans from Arabic) (1)
                2008-11-19 08:05:51 : Forum love/gram (cybertrans from Arabic) (1)
 116
 116.
                2008-11-19 08:06:52 : (referer) forum love/gram (cybertrans from Arabic) (1)
 116.
                2008-11-19 15:01:00: The hills jihadist without inflicting (cybertrans from Arabic) (10)
 116
                2008-11-19 15:14:13 : (referer) the hills jihadist without inflicting (cybertrans from Arabic) (6)
 116.
                2008-11-19 15:33:19 : Waziristan (cybertrans from Arabic) (1)
 116.
                2008-11-19 04:24:44 : Scandals (cybertrans from Arabic) (2)
 116
                2008-11-19 04:24:59 : (referer) scandals (cybertrans from Arabic) (1)
                2008-11-19 04:29:29 : News (cybertrans from Arabic) (1)
 116.
 116.
                2008-11-19 04:30:04 : Forum soil (cybertrans from Arabic) (1)
 116
                2008-11-19 04:31:51 : (referer) forum soil (cybertrans from Arabic) (1)
                         </d>

< d∩u ≥</p>
```

**Workflow Values** 

Workflow XML

### X-KEYSCORE SIGDEV



 EX: Targets pass links to videos, use XKS to discover new targets who have viewed those videos

In HB 00215-09, he promises that the newest video will be ready very soon, and then sends these two links:

http://www.load.to/.	
http://www.files.to/get/	

Datetime:	2 Weeks 💌	Start:	2008-12-23	00:00	Stop:	2009-01-06	23:59 🜲	Q
HTTP Type:				~				
Host:	www.files.to							
URL Path:	/get/							

## X-KEYSCORE SIGDEV



Datet	TS 📤	USERID	PHONE	USER_A		ACTIVITY	USER_B
2008	20081231 2246	506Z			<emailaddr></emailaddr>	logged in (email)	59.
	20081231 2249	949Z			≤emailAddr>	logged in (email)	59.
	20081231 2249	949Z			<emailaddr></emailaddr>	logged in (email)	59.
	20081231 2249	949Z			≤emailAddr≥	logged in (email)	59.
	20081231 2249	952Z			<emailaddr></emailaddr>	logged in (email)	59.
	20081231 2249	952Z			≤emailAddr≥	logged in (email)	59.
	20081231 2249	952Z			≤emailAddr≥	logged in (email)	59.
	20081231 2250	018Z			≤emailAddr≥	logged in (email)	59.
	20081231 2250	021Z			≤emailAddr≥	logged in (email)	59.

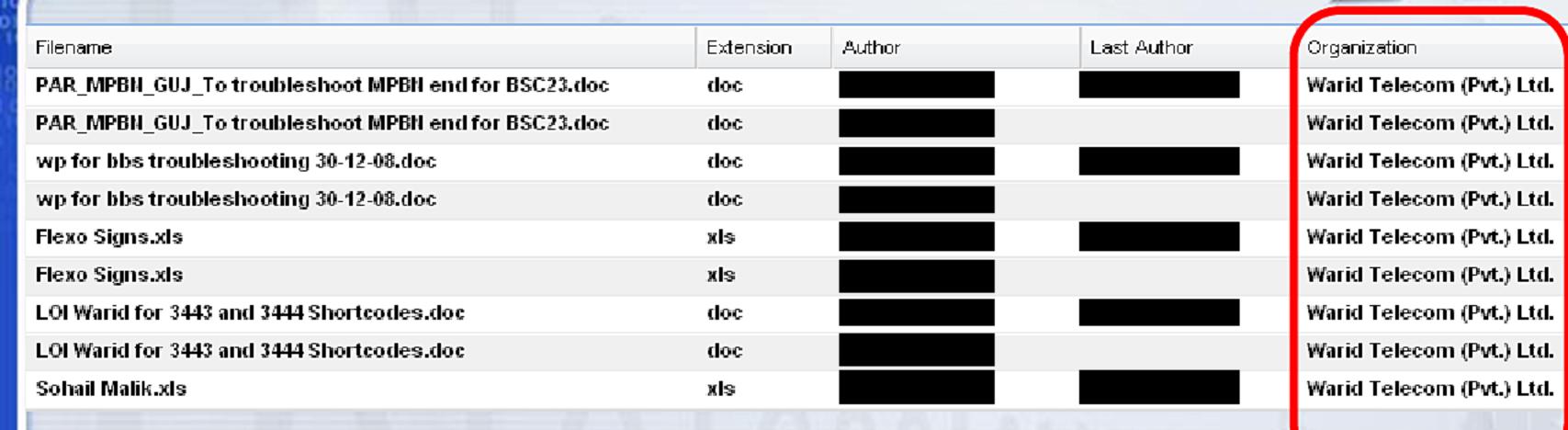
### X-KEYSCORE SIGDEV



How to find technical documents of interest

One Idea: Take advantage of the properties exploited as meta-data by X-KEYSCORE like the Author and Organization

Lets look for all documents where the organization field is the company we're interested in, ex: Warid Telecom



Many of these files may have not been selected, because either there was no strong selector associated or the strong selector(s) weren't tasked for collection



Questions?

@nsa

xkeyscore@nsa

# HTTP Activity



/search?tab=urdu&order=sortboth&q=musharraf&start=3&scope=urdu&link=next HTTP/1.1 Accept: \*/\* Referer: http://search.bbc.co.uk/search?tab=urdu&order=sortboth&q=musharraf&start=2&scope=urdu Accept-Language: en-us Accept-Encoding: gzin. deflate User-Agent Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1) Host: search.bbc.co.uk Cookie: BBC-UID=b479a5f4ad230a53063d513630203acb22684634a0e0b164c45f96efc054cf950Mozilla%2f4%2e0%20%28co Cache-Control: max-stale=U Connection: Keen-Alive X-BlueCoat-Via: 66808702E9A98546 **URL Path URL Args** Host tab=urdu&order=sortboth&q=musharraf&start=3&scope=urdu&link=next search.bbc.co.uk /search Search Terms Language Browser Via Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1). musharraf en

Referer

http://search.bbc.co.uk/search?tab=urdu&order=sortboth&q=musharraf&start=2&scope=urdu

Cookie

BBC-UID=b479a5f4ad230a53063d513630203acb22684634a0e0b164c45f96efc054cf950Mozilla%2f4%2e0%20%28com

TOP SECRET//COMINT//ORCON, REL TO USA, AUS, CAN, GBR and NZL//20291123

# Query Hierarchy



